О П Т И М А Л Ь Н О Е   П Л А Н И Р О В А Н И Е

Сборник трудов
1967 г.   Института математики СО АН СССР   Выпуск 7

А. А. КАПЛАН

НЕКОТОРЫЕ СВОЙСТВА ОПЕРАТОРА ПОКООРДИНАТНОГО СПУСКА

Настоящая статья посвящена исследованию свойств оператора покоординатного спуска применительно к задачам выпуклого программирования. Известно, что метод покоординатного спуска, вообще говоря, не приводит к решению задачи выпуклого программирования, однако, разумное комбинирование его, например, с некоторыми градиентными методами, позволяет получить решение задачи более эффективно, чем при применении градиентных методов в чистом виде.

В случае, когда допустимой областью является все пространство или его часть, задаваемая ограничениями вида $a_i \leq x_i \leq b_i$, при дополнительных условиях относительно гладкости и при строгой выпуклости минимизируемой функции метод покоординатного спуска, как легко показать [I], приводит к искомому минимуму.

Рассматривается задача минимизации выпуклой функции φ на телесном выпуклом ограниченном замкнутом множестве X арифметического n -мерного пространства R^n. Функция φ предполагается заданной на открытом множестве $\tilde{X} \supset X$.

Свойства оператора покоординатного спуска будут изучаться применительно к этой задаче.

Введем некоторые обозначения.

Через A_i обозначим оператор из R^n в R^{n-1}, сопоставляющий точке $x=(x_1, x_2, \ldots, x_n)$ точку $\xi=(\xi_1, \xi_2, \ldots, \xi_{n-1})$ с координатами $\xi_j = x_j$, $j=1,2,\ldots,i-1$, $\xi_j = x_{j+1}$, $j=i,\ldots,n-1$.

Далее:

$$G_i = \{ A_i(x) : x\in X \}, \quad i=1,2,\ldots,n;$$

$$E_i(\xi)=\{x\in X : A_i(x)=\xi\}, \quad i=1,2,\ldots,n;$$

$$M_i(\xi)=\max_{x\in E_i(\xi)} x_i, \quad i=1,2,\ldots,n;$$

$$m_i(\xi)=\min_{x\in E_i(\xi)} x_i, \quad i=1,2,\ldots,n;$$

$$\Omega_i(\xi)=\{x_i : \varphi(x_1,\ldots,x_n)=\min_{y\in E_i(\xi)} \varphi(y)\}, \quad i=1,2,\ldots,n;$$

$$\overline{\Omega}_i(x)=\{y\in X : \varphi(y)=\min_{z\in E_i(A_i(x))} \varphi(z)\}, \quad i=1,2,\ldots,n.$$

Заметим, что лишь в случае строгой выпуклости $\varphi(x)$ можно гарантировать одноточечность всех $\Omega_i(\xi)$ и $\overline{\Omega}_i(x)$.

В силу выпуклости множества X введенные функции M_i и m_i являются, соответственно, вогнутыми и выпуклыми. Следовательно, они непрерывны во внутренних точках множеств G_i. Если предположить, что граница множества X не содержит отрезков, параллельных координатным осям, т.е. для любой граничной точки $\xi\in G_i$ соответствующее множество $E_i(\bar\xi)$ состоит из единственной точки, то рассматриваемые функции M_i и m_i будут непрерывны также и в граничных точках. Непрерывность M_i и m_i в граничных точках непосредственно следует из того, что при сделанном предположении, каковы бы ни были граничная точка $\bar\xi\in G_i$ и $\varepsilon>0$, найдется такое $\delta>0$, что расстояние от любой точки множества $\bigcup_{\xi\in P_i} E_i(\xi)$ до $E_i(\bar\xi)$ меньше ε (здесь P_i – параллелепипед $\bar\xi_j-\delta\le\xi_j\le\bar\xi_j+\delta$, $j=1,\ldots,n$, $j\ne i$)

Имеют место приводимые ниже теоремы I и 2.

Т е о р е м а I. Если $\varphi(x)$ - строго выпуклая функция и граница множества X не содержит отрезков, параллельных осям координат, то операторы Q_i непрерывны.

Д о к а з а т е л ь с т в о. Непрерывность функций Q_i в граничных точках $\xi \in G_i$ непосредственно следует из одноточечности множеств $E_i(\xi)$ и непрерывности функций M_i и m_i. Допустим теперь, что при некотором i функция Q_i не является непрерывной в некоторой внутренней точке $\xi^0 \in G_i$. Тогда найдется такая последовательность $\xi^j \in G_i$, сходящаяся к ξ^0, что

$$\lim_{j\to\infty} Q_i(\xi^j) = a \neq Q_i(\xi^0). \qquad (I)$$

Для произвольного $x_i \in [m_i(\xi^j), M_i(\xi^j)]$ имеем

$$\varphi(\xi_1^j,\ldots,\xi_{i-1}^j, Q_i(\xi^j), \xi_i^j,\ldots,\xi_{n-1}^j) \leq \varphi(\xi_1^j,\ldots,\xi_{i-1}^j, x_i, \xi_i^j,\ldots,\xi_{n-1}^j), \qquad (2)$$
$$j=1,2,\ldots$$

Далее, при любых $\delta_1>0$ и $\delta_2>0$ в силу непрерывности M_i и m_i неравенства $M_i(\xi^j) \geq M_i(\xi^0) - \delta_1$ и $m_i(\xi^j) \leq m_i(\xi^0)+\delta_2$ справедливы при достаточно больших j. Так как во внутренней точке ξ^0

$$M_i(\xi^0) > m_i(\xi^0),$$

δ_1 и δ_2 можно выбрать так, что

$$M_i(\xi^0) - \delta_1 > m_i(\xi^0) + \delta_2 .$$

Интегрируя неравенство (2), получаем неравенство

$$\varphi(\xi_1^j,\dots,\xi_{i-1}^j, \bar{Q}_i(\xi^j), \xi_i^j,\dots,\xi_{n-1}^j)\cdot[M_i(\xi^\circ)-m_i(\xi^\circ)-\delta_1-\delta_2]\le$$

$$\le\int\limits_{m_i(\xi^\circ)+\delta_2}^{M_i(\xi^\circ)-\delta_1}\varphi(\xi_1^j,\dots,\xi_{i-1}^j, x_i, \xi_i^j,\dots,\xi_{n-1}^j)\,dx_i,$$

справедливое при достаточно больших j

Переходя в последнем неравенстве к пределу при $\xi^j\to\xi^\circ$, имеем

$$\varphi(\xi_1^\circ,\dots,\xi_{i-1}^\circ, a, \xi_i^\circ,\dots,\xi_{n-1}^\circ)\le$$

$$\le\frac{1}{M_i(\xi^\circ)-m_i(\xi^\circ)-\delta_1-\delta_2}\int\limits_{m_i(\xi^\circ)+\delta_2}^{M_i(\xi^\circ)-\delta_1}\varphi(\xi_1^\circ,\dots,\xi_{i-1}^\circ, x_i, \xi_i^\circ,\dots,\xi_{n-1}^\circ)\,dx_i .$$

Отсюда, ввиду произвола в выборе δ_1 и δ_2,

$$\varphi(\xi_1^\circ,\dots,\xi_{i-1}^\circ, a, \xi_i^\circ,\dots,\xi_{n-1}^\circ)\le\varphi(\xi_1^\circ,\dots,\xi_{i-1}^\circ, x_i, \xi_i^\circ,\dots,\xi_{n-1}^\circ)$$

для любого $x_i\in(m_i(\xi^\circ), M_i(\xi^\circ))$, а по непрерывности φ это верно и для замкнутого промежутка.

Так как φ - строго выпуклая функция, отсюда следует, что $a=\bar{Q}_i(\xi^\circ)$, вопреки предположению (I). Полученное противоречие доказывает утверждение теоремы.

Рассмотрим теперь оператор $\bar{Q}$, представляющий суперпозицию операторов $\bar{Q}_i$:

$$\bar{Q}(x)=\bar{Q}_n(\bar{Q}_{n-1}(\dots(\bar{Q}_1(x))\dots)).$$

Из доказанной теоремы следует **непрерывность этого оператора.**

О п р е д е л е н и е. Назовем точку $x\in X$ квази-неподвижной, если $\bar{Q}(x)=x$.

Приводимый ниже пример показывает, что при сделанных предположениях квази-неподвижная точка не обязательно является точкой минимума функции φ, даже если функция φ непрерывно дифференцируема.

Пример I. Для функции $\varphi(x)=(x_1+2)^2+(x_2-2)^2$, рассматриваемой на множестве $X=\{x: x_1^2+x_2^2\leq 1\}$, точка $x_1=0$, $x_2=1$ является квази-неподвижной, а минимум этой функции достигается в точке $x_1=-\frac{1}{\sqrt{2}}$, $x_2=\frac{1}{\sqrt{2}}$.

Пусть $x^0\in X$ и последовательность $x^0, x^1, \ldots$ удовлетворяет условию $x^i=\bar{Q}(x^{i-1})$. Оставляя в стороне вопрос о единственности предельной точки такой последовательности, докажем следующую теорему.

Теорема 2. Если граница множества X не содержит отрезков, параллельных координатным осям, то любая предельная точка $\bar{x}$ последовательности $x^0, x^1, x^2, \ldots$ является квази-неподвижной[x)].

Доказательство. Последовательность $x^0, x^1, \ldots$ ограниченная. Предположим, что $\{x^{i_k}\}$ - ее подпоследовательность, сходящаяся к пределу $\bar{x}$, причем $\varphi(\bar{Q}(\bar{x}))\neq\varphi(\bar{x})$. Пусть $\varphi(\bar{x})-\varphi(\bar{Q}(\bar{x}))=K>0$. Обозначим

$$\mathfrak{M}=\{x: x\in X,\ \varphi(x)\leq\varphi(\bar{x})-\frac{K}{2}\}.$$

Очевидно, что $\mathfrak{M}$ - выпуклое множество.

В силу непрерывности $\varphi(x)$ и $\bar{Q}(x)$ для всякого $\varepsilon>0$ найдется N такое, что $|\varphi(\bar{Q}(x^{i_j}))-\varphi(\bar{Q}(\bar{x}))|<\varepsilon$ для $j>N$. Если $\varepsilon=\frac{K}{2}$, то

$$\varphi(\bar{Q}(x^{i_j}))\leq\varphi(\bar{Q}(\bar{x}))+\varepsilon=\varphi(\bar{x})+\varepsilon-K\leq\varphi(\bar{x})-\frac{K}{2},$$

x) В работе [I] при доказательстве теоремы, аналогичной теореме 2, допущена неточность.

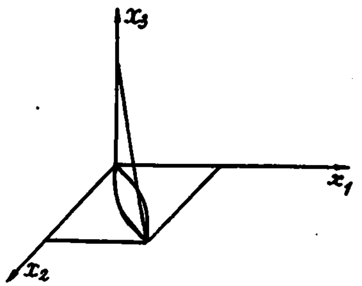
т.е. $\bar{Q}(x^{i_j}) \in \mathfrak{M}$. Отсюда $x^{i_{j+m}} \in \mathfrak{M}$ при любом m вопреки тому, что бесконечное число элементов $x^{i_{j+m}}$ должно принадлежать последовательности $\{x^{i_k}\}$.

Отсюда следует, что $\varphi(Q(\bar{x})) = \varphi(\bar{x})$, а значит $Q(\bar{x}) = \bar{x}$.

Отметим, что предположение относительно границы множества X существенно. Если граница X содержит отрезок, параллельный i -той координатной оси, функции M_i, m_i, Q_i и операторы $\bar{Q}_i$, $\bar{Q}$ могут терпеть разрыв в граничных точках множества G_i . Простой пример сообщен автору Р.Э.Вальским. Приводимый ниже пример показывает, кроме того, что и теорема 2 в этом случае также неверна.

Пример 2. Множество X – конус с вершиной в точке (0,0,I). Основание конуса – выпуклое множество в плоскости $x_3 = 0$, задаваемое неравенствами $(x_1-1)^2 + x_2^2 \leq 1$, $x_1^2 + (x_2-1)^2 \leq 1$, функция $\varphi(x)$ равна

$$x_1^2 + x_2^2 + (x_3-1)^2 .$$



Заметим, что в точках границы множества G_3 , за исключением точки $\xi_1 = 0$, $\xi_2 = 0$ имеет место $M_3 = m_3$. В точке $\bar{\xi} \in G_3$ с координатами $\bar{\xi}_1 = 0, \bar{\xi}_2 = 0$ имеем $M_3(\bar{\xi}) = 1$, $Q_3(\bar{\xi}) = 1$, в любой другой точке на границе G_3 $M_3 = 0$, $Q_3 = 0$.

Следовательно, M_3 и Q_3 терпят разрыв в точке $\bar{\xi}$.

Начиная покоординатный спуск из точки $(\frac{1}{2}, \frac{1}{2}, 0)$, мы получим на участке дуги $x_1^2 + (x_2-1)^2 = 1$, принадлежащем допустимой области, последовательность точек, сходящуюся к $\bar{x} = (0,0,0)$. Однако $Q'(\bar{x}) = (1,0,0)$, т.е. $Q(\bar{x}) \neq \bar{x}$.

В заключение хочу поблагодарить Г.Ш.Рубинштейна за внимание к работе.

ЛИТЕРАТУРА

1. I.Warga. Minimizing certain convex function. I.Industr. and Appl.Math., vol. II, 3, 1963.